

Scene perception guided crowd anomaly detection

Xuguang Zhang^a, Dingxin Ma^a, Hui Yu^{b,*}, Ya Huang^b, Peter Howell^b, Brett Stevens^b

^a Hangzhou Dianzi University, Hangzhou 310018, PR, China

^b University of Portsmouth, Portsmouth PO1 2DJ, UK

ARTICLE INFO

Keywords:

Video surveillance

Crowd anomaly detection

Scene perception

Fluid forces

ABSTRACT

Crowd anomaly detection has been a research hotspot in the field of video surveillance in recent years. In most existing methods, the accuracy of anomaly detection dominantly relies on the acquisition of regions of interest (ROI) and feature extraction. However, the randomness of ROI segmentation and crowd group selection usually cannot guarantee a robust performance and thus may lead to false detection sometimes. To address these issues, this paper proposes a scene perception-based approach combining the fluid forces expression and psychological theory. The proposed method firstly introduces a flow field visualization technology called line integral convolution to segment the moving pedestrians in the scene. Then, a scene perception-guided clustering strategy is proposed to cluster the consistency crowd group. Scene perception strategy is in line with the psychological criteria of human cognition. In clustering, it makes more reasonable use of various attributes of the pedestrians. To ensure a robust detection of the pedestrian group, we propose a fluid feature concept which considers both mass force and surface force. For each consistency group, two types of features including the image appearance feature and fluid feature are combined to describe pedestrian motion. The experimental results show that the proposed method achieves higher accuracy in comparison with some existing methods in terms of both frame-level and pixel-level measurements.

1. Introduction

Crowd video surveillance plays an important role in the field of public safety management. It usually includes tasks such as crowd analysis [1], crowd counting [2] and crowd anomaly detection [3]. Among those, the detection of abnormal behavior is more directly related to public safety. Crowd anomaly behavior can be generally divided into two types including local anomaly and global anomaly. (1) Local anomaly refers to the behavior of some pedestrians that is different from other pedestrians around them, such as the scenarios of vehicle entry and pedestrian retrograde. (2) Global anomaly means that the behavior of a crowd in the scene is abnormal, such as collective aggregation and escape. Compared with the global anomaly, local anomaly has a higher probability of occurrence and is more difficult to detect. In addition, global anomaly is usually formed by the diffusion of local anomaly. If the initial local anomaly is accurately detected, subsequent global anomaly can be avoided in many cases. Generally, there are two kinds of methods for crowd local anomaly detection. One is the trajectory-based method [4-10], the other is features-based method [11-25]. Trajectory-based method use target tracking algorithms to obtain the trajectories of each pedestrian.



Fig. 1. The division of the ROI. (a) Traditional ROI; (b) Proposed consistency group

Anomalies are then detected by learning the potential knowledge of trajectories in normal situations. However, such kind of methods is only suitable for a small-scale crowds. This is because there are inevitable overlaps and occlusions when the crowd density increase, so it is difficult to obtain an accurate trajectory, which will seriously reduce the accuracy of anomaly detection.

Features-based methods avoid these problems by extracting appearance or motion features, such as Histogram of Oriented Gradient (HOG) [26] and Histogram of Optical Flow (HOF) [27], to learn activity patterns of the target. Usually, video frames are divided into multiple regions of interest (ROI) in advance. The ROI usually includes some 2-D image boxes or 3-D video boxes using 3D tracking [28], which are non-overlapping and have the same size. As shown in Fig. 1(a), because ROIs are randomly segmented, some boxes like the red box contain both normal pedestrians and abnormal vehicles, which leads to inaccurate feature extraction. In addition, in some

* Corresponding author.

E-mail address: hui.yu@port.ac.uk (H. Yu)

boxes e.g. the blue box in Fig. 1 (a), the pedestrian integrity is undermined because their bodies are divided into many parts leading to wrong feature extraction.

To address the above issues, we propose an anomaly behavior detection method based on scene perception combining the fluid forces expression theory. In practice, it is not easy to cluster pedestrians with the same movement pattern into a group, because pedestrians have many attributes, such as position, speed and direction of movement. The scene perception model in the field of psychology provides us with a mechanism to deal with multiple attributes of objects [29]. The pattern of processed human spatial and attribute information in the scene is consistent with the separation processing mechanism. Postma et al. [30] found that the processing of spatial information was earlier than, and independent of, the attribute information through the task of reproduction. In the perceptual scene, the preprocessing of spatial information can provide a good foundation for us to understand the world that we see, more effectively. And the subsequent processing of attribute information promotes the basic understanding to complete cognition. This provides us a basis for the clustering of pedestrian groups. Based on this idea, we have developed a divide and conquer clustering strategy to cluster the motion crowd into consistency groups and extracted features for each pedestrian group. The consistency group is shown in Fig. 1(b). The framework of the proposed method is shown in Fig. 2.

More specifically, we integrate the image appearance and fluid forces features to describe the crowd motion in a consistency group. The image appearance features are calculated using a gray-level co-occurrence matrix (GLMC). The fluid forces features are described by calculating mass force and surface force. The fluid forces model was first proposed to study the motion characteristics of fluid particles in Smooth Particle Hydrodynamics (SPH) [31]. In this paper we adapted the fluid forces model as a feature descriptor to detect crowd anomaly motion. This is based on the assumption that the movement of different pedestrians lead to the changes of forces between particles inside. Our contribution can be summarized as follows:

- We developed a novel consistency group method to overcome the drawback of the traditional ROI method. The consistency group only contains complete pedestrians with similar spatial and attribute information, which protects the integrity and consistency of the group.
- Based on scene perception, a divide and conquer clustering strategy is proposed to cluster the motion crowd. The motion crowd is then divided into multiple consistency groups according to the spatial and attribute information, which can use the attributes of the pedestrians more reasonably, such as position (spatial information), speed and direction (attributes information).
- A fluid descriptor is introduced based on the analysis of a fluid forces model. The fluid forces features are extracted by calculating the mass force and surface force of particles in space-time, which can better characterize crowd motions to distinguish abnormal and normal events in an efficient way.

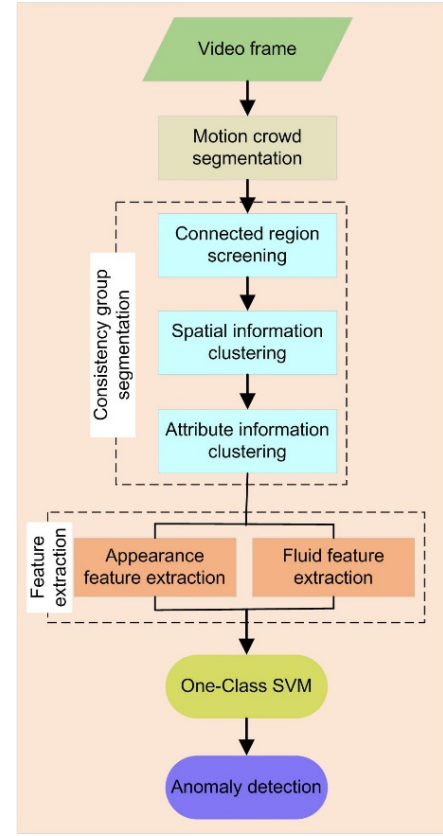


Fig. 2. The framework of the proposed method

The remainder of this paper is organized as follows: an overview of related work is provided in Section 2. Section 3 introduces the method of moving pedestrian segmentation. In Section 4, a divide and conquer clustering strategy is presented to turn a crowd into consistency groups based on scene perception. Section 5 describes the methods of image appearance and fluid features' extraction. Section 6 describes the anomaly detection and localization. Detailed experimental results and discussions are given in Section 7. Finally, Section 8 concludes this paper.

2. Related work

The methods of crowd anomaly detection are mainly divided into two categories: trajectory-based methods and features-based methods. In the first category, the methods obtain the trajectories of moving pedestrians through applying accurate target tracking algorithms. By learning the potential knowledge of trajectories in normal situations, anomalies can be detected. Zhou et al. [6] used KLT to track the trajectory of moving objects. And then abnormal trajectories are recognized through hidden Markov model (HMM). Anjum et al. [7] transformed the trajectories into multiple feature spaces, which were then clustered into different categories for anomaly detection. Spatio-temporal information was also used to detect and locate abnormal behavior of complex targets [8, 32]. Zhang et al. [9] tracked crowd points to capture spatial and temporal information to present a novel feature called Point Trajectory-based

Histogram of Optical Flow (PT-HOF) for anomaly detection. Dou et al. [10] calculated the optical flow field of corresponding joint points and used their trajectories to remove noise. Then Support Vector Machine (SVM) was used to determine abnormal behavior. Although trajectory-based features are high-level semantic, such methods are only suitable for small-scale scenes. With the increase of density of a crowd, there are inevitable overlaps and occlusions. As a result, the accuracy of anomaly detection will seriously reduce. As for the second category, video frames are usually divided into multiple regions of interest (ROI) in advance. Then features from ROIs will be extracted to learn activity patterns. Hao et al. [11] proposed a spatial-temporal texture (STT) extraction algorithm, which can effectively extract video textures with abundant crowd motion details. Then gray level co-occurrence matrix (GLCM) is formulated to detect anomalies. Kaltsa et al. [12] designed the histograms of oriented swarms (HOS) based on the application of swarm intelligence. Then HOS and the well-known histograms of oriented gradients (HOG) were combined to build a descriptor to represent the motion and appearance information. Li et al. [13] used 3D gradient maps and optical flow maps to build a two-stream framework for anomaly detecting. Li et al. [14] proposed an effective feature histogram variance of optical flow angle (HVOFA) and combined it with motion energy to find abnormal events. Khan et al. [15] obtained the neighboring information from surrounding super pixels to create a feature. Then univariate Gaussian discriminant analysis and a K-means algorithm were used to classify abnormal situations. The effectiveness of this type of method depends on the selection of features.

Scene perception focuses on how people perceive and process complex real-world information [33]. In real-life, information obtained by observing an object usually includes both spatial and attribute information. In previous studies, researchers have demonstrated that these two kinds of information are processed separately [34]. In addition, the processing of spatial information is earlier than attribute information [35]. Jiang et al. [36] found that in short-term visual memory, the spatial structure of the scene was first formed based on location information of object. Then the color, shape and other attribute information of the object were further processed and coded based on spatial structure. The spatial structure is preferentially processed in visual short-term memory, which plays a vital role.

Crowd movement has strong physical properties, so many works have applied physical models to analyze crowd videos [37]. Based on the physics properties, the methods of crowd behavior analysis can be expressed as fluid dynamics [38], energy, entropy models [39] and complex system science [40]. Among them, the fluid dynamics model has been widely used. Since crowd motion is similar to fluid flow, many contributions consider crowd motion as a fluid flow to analyze crowd movement. Some methods have been proposed based on the fluid field analysis. Basavaraj et al. [41] used streak-lines to represent crowd flow to classify the abnormal activity for crowd surveillance. Zhang et al. [42] computed the similarity of streak-lines to segment high-density crowds. Other methods are based on the fluid force

model. Zhang et al. [39] calculated the repulsive force based on the

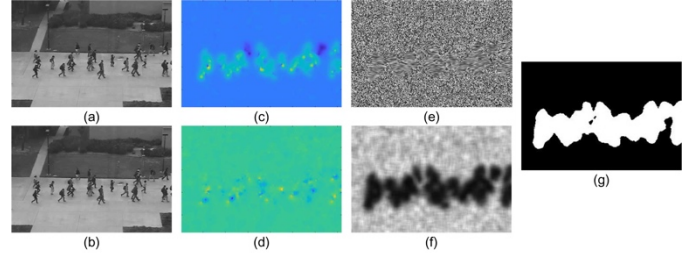


Fig. 3. Moving crowd flow field representation and segmentation. (a) Previous frame; (b) Current frame; (c) Horizontal optical flow; (d) Vertical optical flow; (e) Line integral convolution; (f) Roughness measure; (g) Crowd segmentation.

crowd vector field to construct a weighted network for salient crowd motion detection. Smoothed Particle Hydrodynamics (SPH), which was originally proposed by Gingold [31], is a meshless method for studying the motion characteristics of fluid particles. It has then been gradually applied to many diverse fields [43]. Crowd movement is often described as the flow of fluid, so Ullah et al. [44] applied SPH to the study of crowd consistent movement and proposed a Density Independent Hydrodynamics Model (DIHM) to detect crowd coherency motion.

3. Motion crowd segmentation

In order to judge anomaly behavior in a crowd, the moving targets should be segmented from the background region. Due to background interference, it is very hard to obtain the complete detection of a moving crowd using traditional methods, such as background subtraction and optical flow. Thus, longer spatiotemporal information should be used to help distinguish between moving objects and the background. We also apply the flow field texture visualization technology, which is an effective method to transform flow field data into texture image.

Before visualizing the flow field, it is necessary to obtain the vector field of the crowd's motion. Optical flow method is a classic and effective method to express the pixel motion as a vector field. In this paper, Horn-Schunck's [45] optical flow method is used to obtain the crowd flow field $O = \{(x, y, u, v)\}$. In our previous research [46], we employed a method to obtain the motion region using line integral convolution (LIC). LIC uses the texture synthesis technique to deal with a flow field. During the operation of the LIC method, a low pass filter is used to convolve the input noise texture image along the streamlines calculated from a flow field. Using the LIC method, the crowd flow field can be represented as a texture image. The moving pedestrians can be represented as smooth texture, and the background region can be represented as rough texture. After getting the texture image, a roughness measurement method is used to distinguish the motion region and background region. More detail of the motion crowd segmentation using LIC can be find in [46]. Fig. 3 presents the result of motion crowd segmentation. Fig. 3(a) and (b) show two consecutive frames. The results of optical flow are shown in Fig. 3(c) and (d). The texture visualization can be seen in Fig. 3(e). The difference between the smooth texture and the roughness texture

regions can be represented as Fig. 3(f). In Fig. 3(g), we give the result of moving crowd segmentation.

4. Scene perception based consistency group segmentation

Using LIC based flow field visualization we can segment the moving object. However, it is not able to distinguish the pedestrians moving in different patterns. The moving targets need to be clustered based on a scene perception theory. Inspired by this scene perception theory, we propose a divide and conquer clustering strategy for crowd consistency group detection. As shown in Fig. 4, after segmenting the motion crowd, the connected regions are further screened into small connected regions and big connected regions. Because spatial information is processed in advance in scene perception, the big connected regions are previously clustered to get the spatial proximity groups according to the spatial information. Then, attribute information is processed, so spatial proximity groups and small connected regions are further clustered according to the attribute information. Finally, the motion crowd is divided into multiple consistency groups. The divide and conquer approach avoids the weight distribution problem of various different features in similarity calculation by clustering spatial and attribute information separately.

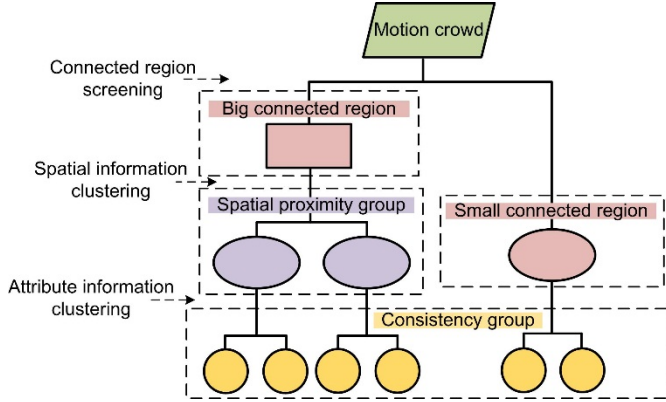


Fig. 4. Divide and conquer clustering strategy

4.1 Divide and conquer clustering strategy

After obtaining the motion crowd, we still cannot directly cluster the motion crowd based on spatial information because some pedestrians will be grouped inappropriately. This will thus destroy the integrity of the group. Because the clustering algorithm only considers the similarity between data, whilst the actual distribution of pedestrians in the scene is ignored.

In order to protect the integrity of the group, small connected regions will not participate in the spatial information clustering. Therefore, we only need to cluster the large connected regions based on spatial information. It can be seen in Fig. 5(c), compared to direct clustering, the spatial consistency group obtained by our method is more reasonable. The screening rule for connected regions is formulated as follows, the connected region with an area larger than η_l will be treated as a big connected region. Otherwise, it is small

connected region.

$$Con = \begin{cases} Con_{big} & \text{if } \sum_{i=1}^M P_i > \eta_l \\ Con_{small} & \text{otherwise} \end{cases} \quad (1)$$

Where P is a pixel in a connected region, $\sum P$ is the number of pixels in the connected region. The threshold η_l is used to distinguish large group such as several pedestrians or vehicle. For a different surveillance scene, the size of the pedestrian or vehicle is related to the distance between the camera and the pedestrian. So, this parameter is selected by experience for different scenes. In addition, each step of clustering is based on the Mean-Shift clustering algorithm. The Mean shift algorithm is a density-based nonparametric clustering algorithm. The basic idea of the algorithm is to find the direction of density increase within a given bandwidth and then move in this direction until convergence [47]. Given n data points $X_i, i=1, 2, \dots, n$, the kernel density estimator computed in the point X is given by:

$$\hat{f}_k(X) = \frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{d(X, X_i)}{h}\right) \quad (2)$$

and the mean shift is:

$$m_{h,G}(X) = \frac{\sum_{i=1}^n X_i g\left(\left\|\frac{d(X, X_i)}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{d(X, X_i)}{h}\right\|^2\right)} - X \quad (3)$$

where $g(x) = -k'(x)$, $k(x)$ is the section function of the kernel function $K(x)$, h is the bandwidth.

For spatial information clustering and attribute information clustering, the bandwidth is set to h_1 and h_2 . We also use Euclidean distance to assess the similarity between pedestrians, the similarity is respectively calculated as:

$$\begin{aligned} d_1(X, X_i) &= \sqrt{(x - x_i)^2 + (y - y_i)^2} \\ d_2(X, X_i) &= \sqrt{(u - u_i)^2 + (v - v_i)^2} \end{aligned} \quad (4)$$

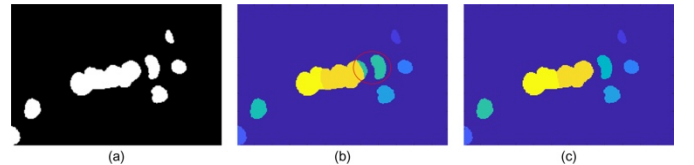


Fig. 5. The comparison of direct clustering with proposed method. (a) Motion crowd; (b) Direct spatial information clustering; (c) Proposed spatial information clustering.

4.2 Noise group merging

The clustering results are usually disturbed by noise when using the scene perception strategy. To solve this issue, we have proposed a noise group merging method. Through observation and analysis, we can judge the noise groups based on the following two assumptions: (1) There are differences between the four limbs' movement and trunk movement of pedestrians. So they are often separately clustered into noise groups. These noise groups can be filtered by setting a threshold η_2 , because their area is much smaller than normal clustering groups. (2) Pedestrian movements are often continuous. So, the motion group detected by individual frames should be noise. These two assumptions

are formulated as:

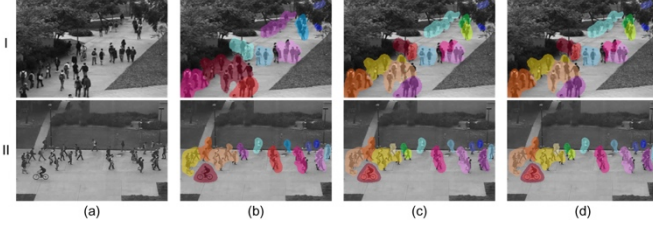


Fig. 6. The result of consistency group segmentation. (a) Input video frame; (b) Spatial information clustering; (c) Attribute information clustering; (d) Noise group merging.

$$G = \begin{cases} G_{noise} & \text{if } \sum_{i=1}^M P_i < \eta_2 \text{ or } \bigcap_{T=t-1}^{t+1} G_T = 0 \\ G_{normal} & \text{otherwise} \end{cases} \quad (5)$$

Where G_{noise} is the noise group, G_{normal} is the normal clustering group. When the group is judged as a noise group, it's merged with the nearest normal clustering group, which is formulated as:

$$G_{normal} = \begin{cases} G_{noise} \cup G_{normal} & \text{if } d(G_{noise}, G_{normal}) = d_{min} \\ G_{normal} & \text{otherwise} \end{cases} \quad (6)$$

Where $d(G_{noise}, G_{normal})$ is the distance between G_{noise} and G_{normal} , d_{min} is the nearest distance from G_{noise} to all G_{normal} . The threshold η_2 is used to distinguish the pedestrians and limbs. We also need to select different thresholds in different scenarios based on experience.

The result of pedestrian clustering is shown in Fig. 6. It worth noting, sometimes, the peripheral areas of the consistency group will gather into a single category. This is because there is a lower optical flow value in the peripheral area. However, such areas will be divided into normal categories during anomaly detection, so the results of anomaly detection will not be affected.

5. Feature expression in consistency groups

After obtaining the consistency groups, the appearance and fluid features of each consistency groups are extracted to detect the anomaly behavior. The extraction for appearance and fluid features will be discussed in detail in the section.

5.1 Appearance feature extraction

Because the abnormal motion areas usually include vehicles, bicycles, etc. there will be some differences in image texture between normal and abnormal groups. In order to extract the appearance features from each consistency group, the texture features are used in our work. Gray-level co-occurrence matrix (GLCM) is an effective method to describe image texture [48]. Some statistics extracted from co-occurrence matrix can be used to describe the texture features of crowd consistency groups.

In GLMC, the probability P_{ij} of each pair of points are calculated as:

$$P_{ij} = \frac{W_{ij}}{n} \quad (7)$$

where W_{ij} is the number of point pairs with the gray level i and gray

level j , n is the total number of point pairs. Based on the probability,

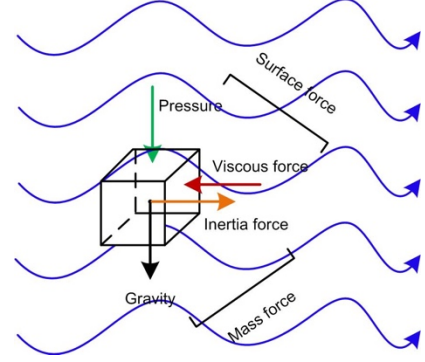


Fig. 7. The model of mass force and surface force

the energy, entropy and contrast of a co-occurrence matrix can be calculated. Energy reflects the uniformity of image gray distribution and the roughness of texture. The energy of a co-occurrence matrix can be calculated as:

$$E = \sum_{i=1}^K \sum_{j=1}^K P_{ij} \quad (8)$$

Entropy reflects the heterogeneity or complexity of an image texture. When all the values in the GLMC are equal or the pixel values show the greatest randomness, the entropy is the greatest. The entropy of a co-occurrence matrix can be calculated as:

$$H = -\sum_{i=1}^K \sum_{j=1}^K P_{ij} \log_2 P_{ij} \quad (9)$$

Contrast reflects the clarity of the image and the depth of texture grooves. The greater the contrast, the deeper the grooves and the clearer the effect. Contrast of a co-occurrence matrix can be calculated as:

$$I = \sum_{i=1}^K \sum_{j=1}^K (i-j)^2 P_{ij} \quad (10)$$

In this paper, four GLCM are constructed, where distance is 1 and respective angle is 0° , 45° , 90° and 135° . So three statistics should be averaged as the final appearance features $\{E, H, I\}$.

5.2 Fluid feature extraction

In this paper, we model the crowd flow as a group of interacting particles. The properties of each particle can be described by physical quantities such as pressure, density and speed. Furthermore, the flow field dynamics model among particles is constructed to express the behavior of crowd groups. The proposed fluid feature is based on Smoothed Particle Hydrodynamics (SPH) [31], which has been widely used in many research fields, such as astrophysics, shock explosion and hydrodynamics. In addition, the SPH model has shown good performance in detecting the consistency of crowd movement [44].

We analyze the forces in crowd motion by considering the Navier-Stokes equation, which is formulated as follows:

$$\rho \left(\frac{\partial v}{\partial t} + v \cdot \nabla v \right) = F_{mas} + F_{sur} \quad (11)$$

According to the N-S equation, we know that fluid motion is mainly affected by the mass force and surface force, the model is shown in Fig. 7. The complete motion pattern of the particles is formed by the combination of these two kinds of forces.

5.2.1 Mass force

Mass force acts on the particles themselves and is proportional to the mass. It is the embodiment of the particle's own motion pattern under the gravitational field and the inertial field. Particles in different motion modes are affected by different mass forces. Mass force is composed of gravity and inertia force, which is mainly determined by the acceleration of the particle. Under the gravitational field, the acceleration of the particle is g . And the inertial acceleration under the inertial field reflects the change of velocity in time, which is formulated as:

$$a = \frac{v_t - v_{t-1}}{T} \quad (12)$$

Where v_t is the speed of current frame, v_{t-1} is the speed of previous frame, T is the duration of a frame. So the mass force can be calculated by:

$$F_{mas} = \sum_{i=1}^N m_i (g + |a|) \cdot K(r_c - r_i, \lambda) \quad (13)$$

Where N is the number of particles in support domain, m is the mass of particles, $K(r_c - r_i, \lambda)$ is the smooth kernel function. It can be understood as a weight function of the extent to which other particles affect the study particle over a range of smooth length λ . The smooth kernel function is non-negative and monotonically decreases as the distance from the center particle increases. A Gaussian kernel function is used in this paper.

For each consistency group, the mass forces should be weighted and averaged based on the distance from the particles in the group to the cluster center particle.

$$\bar{F}_{mas} = \frac{1}{M} \sum_{i=1}^M \frac{1}{d(P_i, G)} F_{mas} \quad (14)$$

Where $d(P_i, G)$ is the distance between particle P_i and consistency group G .

5.2.2 Surface force

Surface force acts on the surface of particles and is related to a superficial area. It is generated by the interaction between particles. If two particles are in different motion modes, the surface forces between them will change significantly. Surface force is mainly divided into two parts: pressure and viscous force. Among them, the pressure reflects the collision between particles. In order to calculate it, the crowd flow is regarded as an ideal fluid and Bernoulli equation is considered, which is formulated as:

$$p + \frac{1}{2} m v^2 + mgh = C \quad (15)$$

Where p is pressure, v is velocity, C is a constant. Since the pedestrians are walking on the ground, the gravitational potential

energy mgh is set as 0. So, pressure is formulated as:

$$F_{pre} = \sum_{i=1}^N (C - \frac{1}{2} m_i v_c^2) S \cdot K(r_c - r_i, \lambda) \quad (16)$$

The viscous force is produced by the velocity difference, which reflects the mutual friction between particles. It is formulated as:

$$F_{vis} = \sum_{i=1}^N \mu S \frac{v_c - v_i}{R} \cdot K(r_c - r_i, \lambda) \quad (17)$$

Where μ is fluid viscosity, S is contact area between particles, v_c is velocity of central particle, v_i is velocity of surrounding particle, R is distance between two particles. The kernel function is set as the Gaussian kernel function. So the surface force can be calculated by:

$$F_{sur} = \sum_{i=1}^N (C - \frac{1}{2} m_i v_c^2 + \mu \frac{v_c - v_i}{R}) S \cdot K(r_c - r_i, \lambda) \quad (18)$$

For each consistency group, the surface force is weighted and averaged as follow:

$$\bar{F}_{sur} = \frac{1}{M} \sum_{i=1}^M \frac{1}{d(P_i, G)} F_{sur} \quad (19)$$

6. Crowd anomaly detection

In our work, the appearance and fluid features are combined as the final descriptor $\{E, H, I, F_{mas}, \bar{F}_{sur}\}$ and one-class SVM (OC-SVM) is used to detect crowd anomaly motion. For anomaly detection, the probability of abnormal events is low, so it is not easy to get many samples in the training set. Just like the UCSD dataset used in our paper, its training set only contains normal events. OC-SVM has good applicability in anomaly detection because its low requirements for training samples. Therefore, we choose OC-SVM as the classifier due to the imbalance of samples. Then, OC-SVM is used to classify the normal and abnormal behaviors of the population, and the calculation cost of OC-SVM is also low.

OC-SVM is developed based on traditional SVM. The main idea of OC-SVM is to treat the origin as an abnormal sample point and other input samples as normal samples. Then the input samples are mapped to high-dimensional space by kernel function, so that there is a hyperplane separating the origin from the input samples and the distance between hyperplane and origin is maximum. In order to find this maximal hyperplane, the basic problem is transformed into the following quadratic programming problem:

$$\begin{aligned} \min_{\omega, \rho} & \frac{1}{2} \|\omega\|^2 + \frac{1}{vN} \sum_{i=1}^N \xi_i - \rho \\ \text{s.t.} & \begin{cases} \omega^T \cdot \phi(x_i) \geq \rho - \xi_i \\ \xi_i \geq 0, i=1, 2, \dots, N \end{cases} \end{aligned} \quad (20)$$

Where x_i is the training data, $i=1, 2, \dots, N$ is the number of training data, $\phi(\cdot)$ is a mapping function which maps primitive space to high-dimensional feature space, ω is the normal vector for hyperplane, ρ is the distance from the hyperplane to the origin, $v \in [0, 1]$ is the upper bound of outliers, ξ_i is the relaxation variable.

Table 1

The setting of variable parameters

Dataset	UCSD Ped1	UCSD Ped2	UMN
Threshold η_1	1800	3300	3280
Threshold η_2	65	140	120
Bandwidth h_1	28	35	40
Bandwidth h_2	0.7	0.9	1.6

Table 2

Comparison of frame-level and pixel-level EER and AUC in ped1

Method	Frame Level		Pixel Level	
	EER	AUC	EER	AUC
Adam [16]	38%	0.649	76%	0.13
SF [17]	31%	0.675	79%	0.197
MPPCA [18]	40%	0.59	81%	0.205
MDT [19]	25%	0.818	58%	0.441
H-MDT [20]	17.8%	-	-	-
STC [21]	15%	-	27%	-
SDL [22]	14.9%	0.932	-	0.716
SCG-SF [51]	16.2%	0.909	-	-
Ours	14%	0.90	26%	0.71

After training, we apply the final descriptor to the OC-SVM model. Because the appearance and fluid features are extracted from each consistency group, an anomaly can be well located as long as the consistency group is judged to be the anomaly group. The following experiments demonstrate that our method accurately localizes different anomalies in many videos.

7. Experimental results

In this paper, two public datasets (UCSD dataset [49] and UMN dataset [50]) are used to test the performance of the proposed method. Then, based on two different levels of measurement (frame-level and pixel-level), we have compared the proposed method with some existing methods for local and global crowd anomaly detection. The experiment was conducted on a computer with a specification of 1.60GHz CPU and 8GB memory.

7.1 Parameter setting

In the proposed method, the parameters including variable parameters and fixed parameters are set for different scenes. Two kinds of variable parameters will be affected by the surveillance scene in this paper. The first one is an area threshold, which is used to screen a big connected region (η_1) and detect the noise group (η_2). The second one is the bandwidth of the divide and conquer clustering, bandwidth of spatial information clustering (h_1) and bandwidth of attribute information clustering (h_2). Because these parameters are related to the size, space distance or moving speed of pedestrians in the image, different parameters need to be selected in different scenes. The specific settings of parameters are given in Table 1. In addition,

the fixed parameters in calculation of fluid features are set as: the

Table 3

Comparison of frame-level and pixel-level EER and AUC in ped2

Method	Frame Level		Pixel Level	
	EER	AUC	EER	AUC
Adam [16]	42%	0.63	76%	0.13
SF [17]	42%	0.556	79%	0.17
MPPCA [18]	30%	0.693	82%	0.13
MDT [19]	25%	0.829	55%	0.42
H-MDT [20]	18.5%	-	-	-
STC [21]	13%	0.92	26%	-
SDL [22]	-	-	-	-
SCG-SF [51]	17.3%	0.902	-	-
Ours	12%	0.93	24%	0.75

smooth length is set as $\lambda = 3$. The mass of each particle is set as $m = 1$. The Bernoulli equation constant is set as $C = 10$. The contact area between particles is set as $S = 1$.

7.2 Evaluation criterions

The Receiver Operating Characteristic (ROC) curve is plotted to measure the detection accuracy. ROC curve is determined by two coefficients, TPR (True Positive Rate) and FPR (False Positive Rate), which can be computed as follows:

$$\begin{aligned} TPR &= \frac{TP}{TP + FN} \\ FPR &= \frac{FP}{FP + TN} \end{aligned} \quad (21)$$

Then Area Under Curve (AUC) and Equal Error Rate (EER) are used to quantify the ROC performance.

7.3 Crowd anomaly detection for local anomaly

The UCSD dataset consists of two subsets with different perspectives, ped1 and ped2. Ped1 contains 34 training video sequences and 36 test video sequences with a resolution of 238×158 (only 10 sequences with the ground truth available are used in this experiment). Ped2 contains 16 training video sequences and 12 test video sequences with a resolution of 360×240. All the training sequences merely include normal events, while testing sequences contain both normal events and anomalies. The normal events in the dataset are just pedestrians on walkways. The anomalies include cyclists, motorists, skaters, wheelchair users etc, which are related to behavior and appearance abnormalities. In order to evaluate our method, two different levels of measurement (frame-level and pixel-level) are considered in our experiments.

(1) Frame-level: As long as one pixel is detected as an anomaly in a test frame, the frame is labeled as an abnormal frame. Although frame-level measurement has been adopted by many researchers, it cannot judge whether the abnormal positioning is correct.

(2) Pixel-level: A frame is considered to be abnormal if at least 40% of the detection overlaps with the ground truth, which emphasizes the correctness of abnormal positioning.

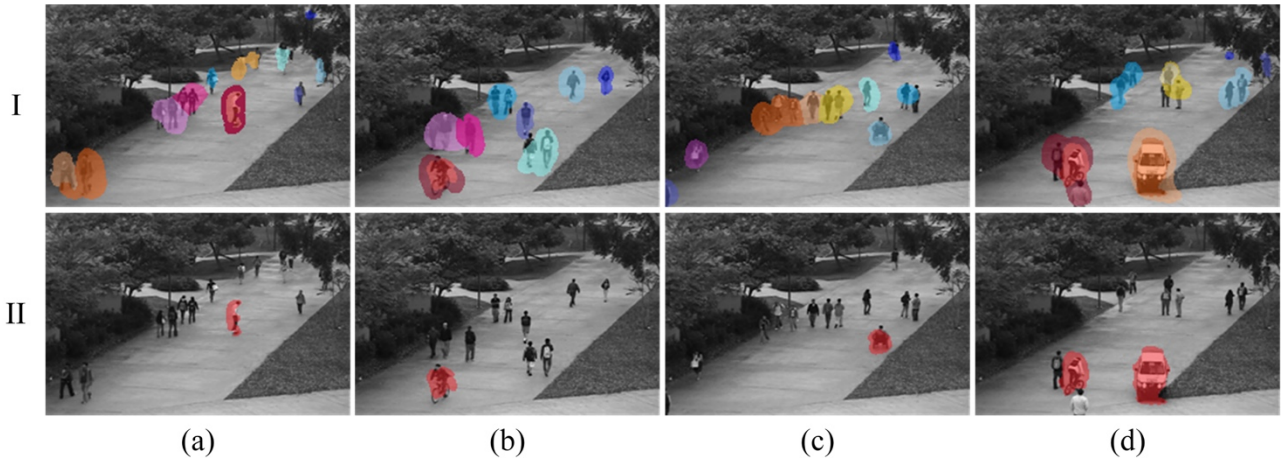


Fig. 8. Sample frame of clustering and anomaly detection results in Ped1. (I) Clustering results. Different consistency groups are represented by different colors; (II) The result of anomalies detection, anomalies are marked in red.

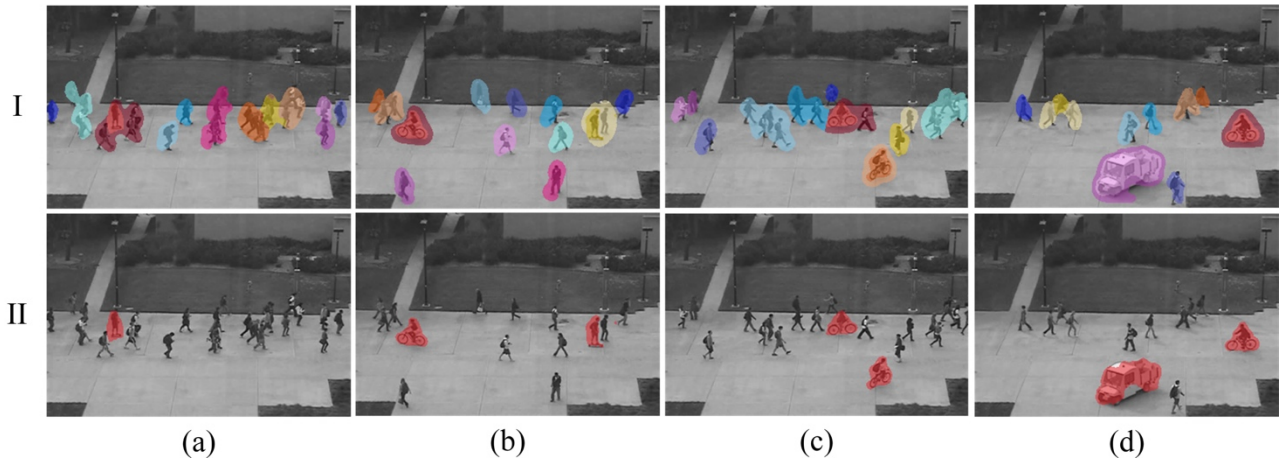


Fig. 9. Sample frame of clustering and anomaly detection results in Ped2. (I) Clustering results. Different consistency groups are represented by different colors; (II) The result of anomalies detection, anomalies are marked in red.

Fig. 8 and Fig. 9 show the anomaly detection results in ped1 and ped2. The top block of the figure depicts the result of motion crowd group clustering in ped1 and ped2. It is shown that pedestrians with different positions and attributes can be well separated. The lower block of the figure depicts some frames that anomalies are successfully detected in ped1 and ped2 respectively. As can be seen in these figures, different kinds of anomalies can be quite well detected, even when multiple anomalies occur in the same frame.

Frame-level and pixel-level ROC curves in ped1 and ped2 are shown in Fig. 10 and Fig. 11, respectively. Based on ROC curves, EER and AUC are computed and listed in Table 2 and Table 3. Because some existing methods did not calculate EER or AUC, while other methods did not use ped2 database. Therefore, we use dashes to represent them in the table.

We have compared the proposed method with some existing methods, which include Adam [16], Social Force (SF) [17], MPPCA [18], Mixture of Dynamic Textures (MDT) [19], H-MDT [20], Spatio-temporal Composition (STC) [21], Structured Dictionary Learning [22] and Sparse Coding Guided Spatiotemporal Feature (SCG-SF) [51]. As shown in Table 2 and 3, our method is competitive in almost both measurements, which is much better than the first five

algorithms. Compared with STC algorithm, the proposed method has a weak advantage. The performance of proposed method is close to the SDL algorithm. In terms of EER index, it is slightly better than SDL method, while in terms of AUC, it is slightly lower than the SDT method. As for SCG-SF method, The proposed method has advantages in ped2 scene. In ped1 scene, the AUC value of SCG-SF method is slightly higher, but the EER value is also larger. Different with AUC, EER is usually used to reflect the error acceptance rate and error rejection rate. From the perspective of algorithm strategy, the reason why the proposed method is competitive is that we use consistency group as the research object. Our method keeps the integrity of the crowd group, so as to improve the accuracy of anomaly detection. In addition, noise was eliminated in the crowd group clustering stage, which reduced false positives in the results.

We calculated the training and testing time of the proposed method. For Ped1 dataset, the training time is about 2062.36 seconds (including feature extraction and classifier training). The average testing time for each frame is about 0.31 seconds. For Ped2 dataset, the training time is about 2007.17 seconds (including feature extraction and classifier training). The average testing time for each frame is 0.86 seconds.

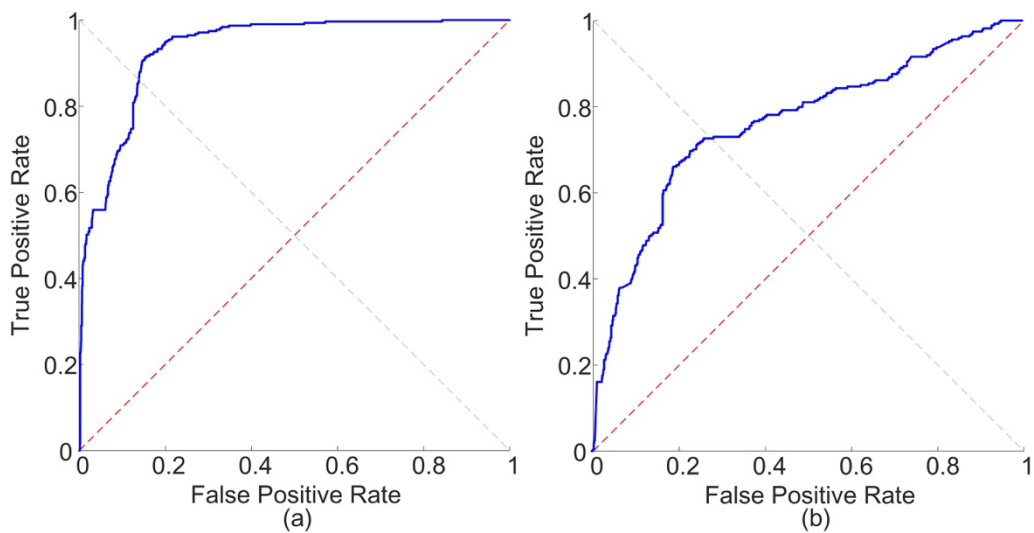


Fig. 10. The ROC curve in UCSD Ped1 dataset. (a) Frame level criterion; (b) Pixel level criterion.

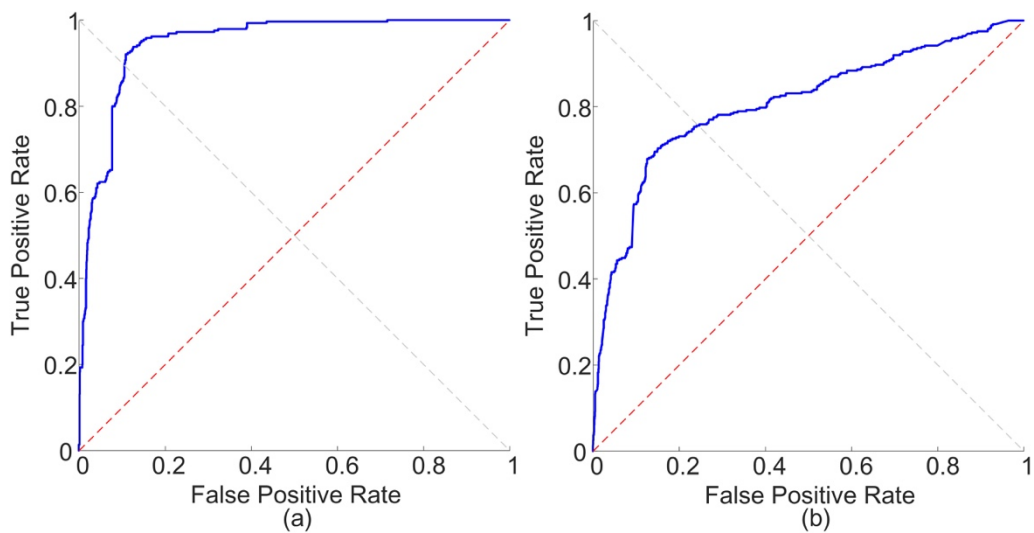


Fig. 11. The ROC curve in UCSD Ped2 dataset. (a) Frame level criterion; (b) Pixel level criterion.

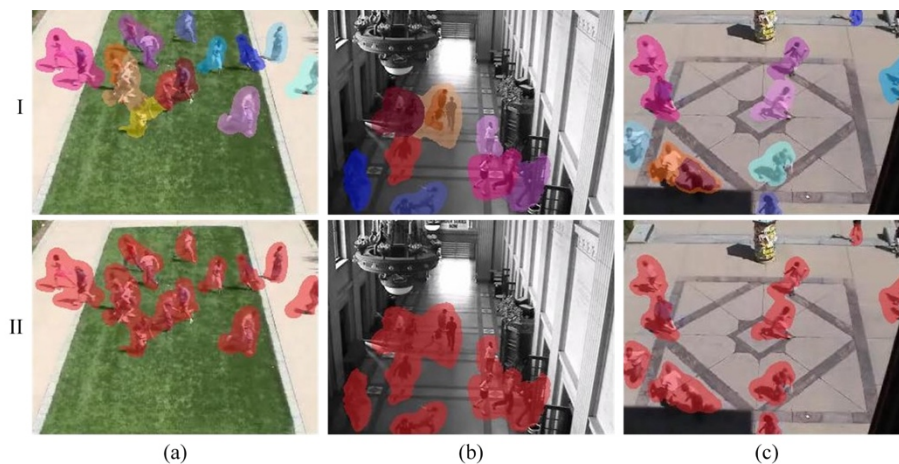


Fig. 12. Sample frames of clustering and anomaly detection results in UMN dataset. (I) Clustering results. Different consistency groups are represented by different colors; (II) Anomalies detection results. Anomalies are marked in red.

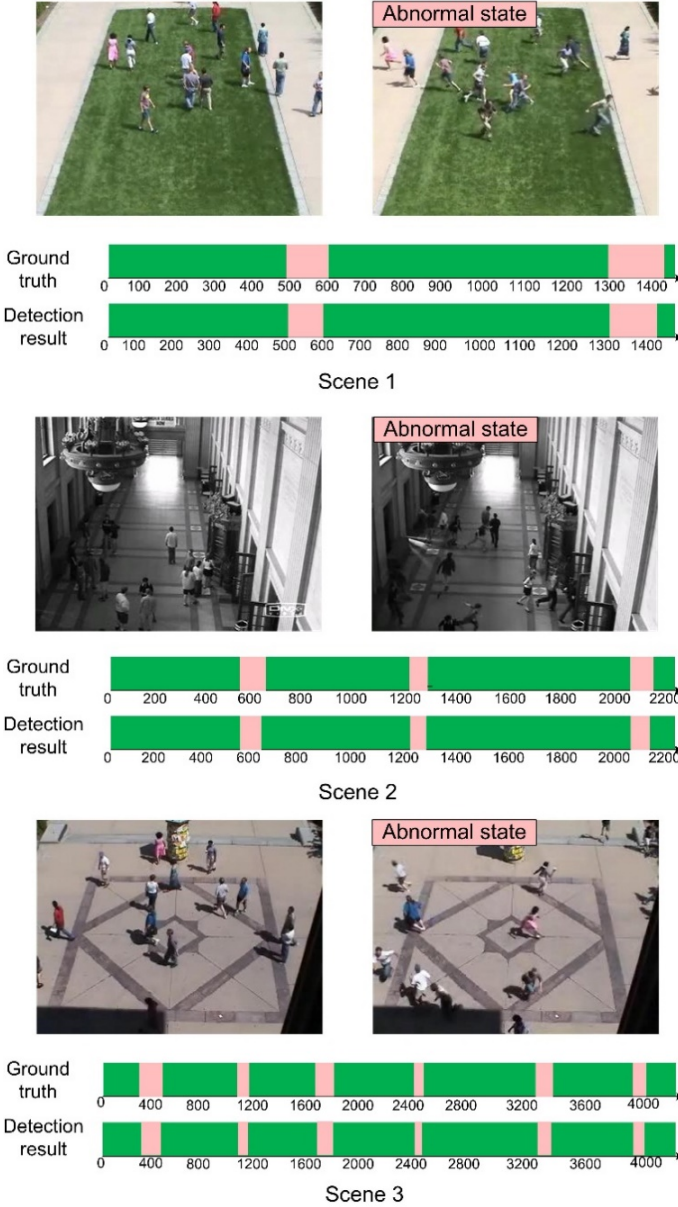


Fig. 13. The qualitative results for global anomalies detection in UMN dataset. The ground truth bar and the detection result bar show the labels of each frame, where green color denotes the normal frames and pink denotes the abnormal frames.

7.4 Crowd anomaly detection for global panic

UMN is a crowd panic behavior dataset, which is comprised of 7739 frames of 320×240 pixels in three different scenes, lawn, indoor and square. In each scene, we select 1000 frames normal behavior for training and the rest frames are used for testing.

The result of motion crowd grouping and anomaly detection are shown in Fig. 12. The first row shows the result of crowd group clustering. And the second row depicts the global anomalies detection result, which shows the effective performance of the proposed method. Fig. 13 shows the qualitative results in three scenes of UMN dataset. The normal/abnormal frames in three scenes are marked as green/pink in the bar.

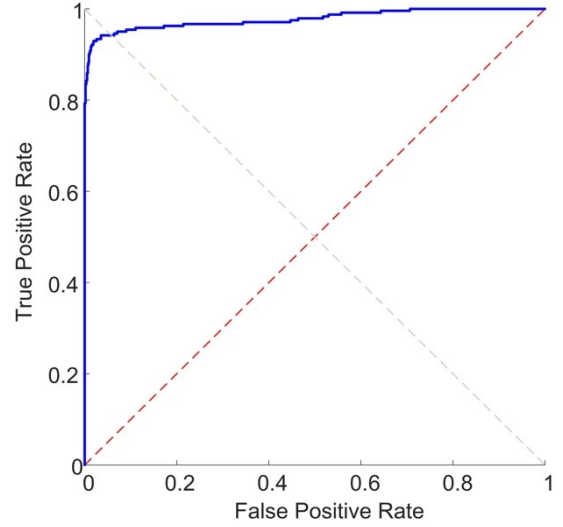


Fig. 14. The ROC curve in UMN dataset

Table 4 Comparison of AUC in UMN dataset

Method	Scene 1	Scene 2	Scene 3	Overall
Optical flow [17]	-	-	-	0.84
SF [17]	-	-	-	0.96
local stat. [23]	-	-	-	0.985
Sparse [24]	0.995	0.975	0.964	-
Chaotic [25]	-	-	-	0.99
H-MDT [20]	-	-	-	0.995
Ours	0.996	0.962	0.987	0.983

Fig. 14 shows the ROC curves. Based on ROC curves, AUC is taken as a quantitative index to compare our method with six other methods: Social Force [17], Optical Flow [17], Local stat. [23], Sparse [24], Chaotic [25] and H-MDT [20]. As shown in Table 4, almost all the methods can achieve high recognition rate for global anomaly detection. The AUC of the proposed method is 0.983, which outperforms optical flow and social force model. We also calculated the AUC of the proposed method in different scenes. For scene 1, the proposed method achieves the best performance, the AUC value is about 0.996. However, the performance of the proposed method in scene 2 is slightly lower than the other two scenes, the AUC value is only about 0.962. For scene 3, our method also performs well, the AUC value is about 0.987. In this experiment, the training time is about 2066.81 seconds (including feature extraction and classifier training). The average testing time for each frame is 0.72 seconds.

8. Conclusion

In this paper, we propose a crowd anomaly behavior detection method based on scene perception and a fluid forces model. Firstly, the motion crowd is segmented by using a flow field visualization based method. Secondly, a divide and conquer clustering strategy is presented to cluster the motion crowd into multiple consistency groups based on a scene perception strategy. Then, by calculating the gray-level co-occurrence matrix of each consistency group, three

statistics (energy, entropy and contrast) are taken as the appearance features. According to the fluid forces model, mass force and surface force are taken as the fluid features. Finally, appearance and fluid features are combined as the final descriptor of each consistency group, and a one-class SVM is used to detect crowd anomalies. The experimental results show that the proposed method can be used to detect crowd global and local anomalies. In the future, we will investigate how to adaptively calculate parameters for specific scenes to improve the adaptability of the method in different cases. And will explore more advanced clustering algorithm [52]. The future work will further explore the psychological model of crowd emotion evolution for feature extraction.

Acknowledgments

This research was supported by National Natural Science Foundation of China (no. 61771418).

References

- [1] Z. Dong, R. Zhang, X. Shao, Y. Li, "Scale-Recursive Network with point supervision for crowd scene analysis," *Neurocomputing*, vol. 384, pp. 314-324, 2020.
- [2] Z. Shi, L. Zhang, Y. Liu, X. Cao, Y. Ye, M. Cheng, G. Zheng, "Crowd Counting with Deep Negative Correlation Learning," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5382-5390, 2018.
- [3] X. Hu, J. Dai, Y. Huang, H. Yang, L. Zhang, W. Chen, G. Yang, D. Zhang, "A weakly supervised framework for abnormal behavior detection and localization in crowded scenes," *Neurocomputing*, vol. 383, pp. 270-281, 2020.
- [4] W. Hu, X. Zhou, W. Li, W. Luo, X. Zhang, S. Maybank, "Active Contour-Based Visual Tracking by Integrating Colors, Shapes, and Motions," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1778-1792, 2013.
- [5] B. T. Morris, M. M. Trivedi, "Trajectory Learning for Activity Understanding: Unsupervised, Multilevel, and Long-Term Adaptive Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2287-2301, 2011.
- [6] S. Zhou, W. Shen, D. Zeng, Z. Zhang, "Unusual Event Detection in Crowded Scenes by Trajectory Analysis," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1300-1304, 2015.
- [7] N. Anjum, A. Cavallaro, "Multifeature Object Trajectory Clustering for Video Analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1555-1564, 2008.
- [8] D. Tran, J. Yuan, D. Forsyth, "Video Event Detection: From Subvolume Localization to Spatiotemporal Path Search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 404-416, 2014.
- [9] K. Zhao, B. Liu, W. Li, N. Yu, Z. Liu, "Anomaly Detection and Localization: A Novel Two-Phase Framework Based on Trajectory-Level Characteristics," *IEEE International Conference on Multimedia & Expo Workshops*, pp. 1-6, 2018.
- [10] Y. Dou, F. Cai, J. Li, C. Wei, "Abnormal Behavior Detection Based on Optical Flow Trajectory of Human Joint Points," *Chinese Control and Decision Conference*, pp. 653-658, 2019.
- [11] Y. Hao, Z. Xu, Y. Liu, J. Wang, J. Fan, "Effective Crowd Anomaly Detection Through Spatio-temporal Texture Analysis," *International Journal of Automation and Computing*, vol. 16, no. 1, pp. 27-39, 2019.
- [12] V. Kaltsa, A. Briassoulis, I. Kompatsiaris, L. J. Hadjileontiadis, M. G. Strintzis, "Swarm Intelligence for Detecting Interesting Events in Crowded Environments," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2153-2166, 2015.
- [13] N. Li, F. Chang, "Video anomaly detection and localization via multivariate gaussian fully convolution adversarial autoencoder," *Neurocomputing*, vol. 369, pp. 92-105, 2019.
- [14] X. Li, W. Li, B. Liu, Q. Liu, N. Yu, "Object-Oriented Anomaly Detection in Surveillance Videos," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1907-1911, 2018.
- [15] M. U. K. Khan, H. Park, C. Kyung, "Rejecting Motion Outliers for Efficient Crowd Anomaly Detection," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 541-556, 2019.
- [16] A. Adam, E. Rivlin, I. Shimshoni, D. Reinitz, "Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 555-560, 2008.
- [17] R. Mehran, A. Oyama, M. Shah, "Abnormal Crowd Behavior Detection using Social Force Model," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 935-942, 2009.
- [18] J. Kim, K. Grauman, "Observe Locally, Infer Globally: A Space-time MRF for Detecting Abnormal Activities with Incremental Updates," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2921-2928, 2009.
- [19] V. Mahadevan, W. Li, V. Bhalodia, N. Vasconcelos, "Anomaly Detection in Crowded Scenes," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1975-1981, 2010.
- [20] W. Li, V. Mahadevan, N. Vasconcelos, "Anomaly Detection and Localization in Crowded Scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18-32, 2014.
- [21] M. J. Roshtkhari, M. D. Levine, "An On-line, Real-time Learning Method for Detecting Anomalies in Videos using Spatio-temporal Compositions," *Computer Vision and Image Understanding*, vol. 117, no. 10, pp. 1436-1452, 2013.
- [22] Y. Yuan, Y. Feng, X. Lu, "Structured Dictionary Learning for Abnormal Event Detection in Crowded Scenes," *Pattern Recognition*, vol. 73, pp. 99-110, 2018.
- [23] V. Saligrama, Z. Chen, "Video Anomaly Detection Based on Local Statistical Aggregates," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2112-2119, 2012.
- [24] Y. Cong, J. Yuan, J. Liu, "Sparse Reconstruction Cost for Abnormal Event Detection," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3449-3456, 2011.
- [25] S. Wu, B. E. Moore, M. Shah, "Chaotic Invariants of Lagrangian Particle Trajectories for Anomaly Detection in Crowded Scenes," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2054-2060, 2010.
- [26] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886-893, 2005.
- [27] N. Dalal, B. Triggs, C. Schmid, "Human Detection using Oriented Histograms of Flow and Appearance," *European Conference on Computer Vision*, vol. 3952, 2006.
- [28] Q. Fu, X. Chen, W. He, "A Survey on 3D Visual Tracking of Multicopters," *International Journal of Automation and Computing*, vol. 16, no. 3, pp. 707-719, 2019.
- [29] J. M. Henderson, A. Hollingworth, "High-level Scene Perception," *Annual Review of Psychology*, vol. 50, no. 1, pp. 243-271, 1999.
- [30] A. Postma, E. H. D. Haan, "What Was Where? Memory for Object Locations," *Quarterly Journal of Experimental Psychology A Human Experimental Psychology*, vol. 49, no. 1, pp. 178-99, 1996.
- [31] R. A. Gingold, J. J. Monaghan, "Smoothed Particle Hydrodynamics: Theory and Application to Non-spherical Stars," *Monthly Notices of the Royal Astronomical Society*, vol. 181, no. 3, pp. 375-389, 1977.
- [32] Y. Hao, Z. Xu, Y. Liu, J. Wang, J. Fan, "Effective Crowd Anomaly Detection Through Spatio-temporal Texture Analysis," *International journal of Automation and Computing*, vol. 16, no. 1, pp. 27-39, 2019.
- [33] J. M. Henderson, "Human Gaze Control During Real-world Scene

Perception," *Trends in Cognitive Sciences*, vol. 7, no. 11, pp. 498-504, 2003.

[34] Mishkin, Mortimer, L. G. Ungerleider, K. A. Macko, "Object Vision and Spatial Vision: Two Cortical Pathways," *Trends in Neurosciences*, vol. 6, no. 10, pp. 414-417, 1983.

[35] Epstein, S. William, "Priming Spatial Layout of Scenes," *Psychological Science*, vol. 8, no. 5, pp. 374-378, 1997.

[36] Y. Jiang, I. R. Olson, M. M. Chun, "Organization of Visual Short-term Memory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 26, no. 3, pp. 683-702, 2000.

[37] X. Zhang, Q. Yu, H. Yu, "Physics Inspired Methods for Crowd Video Surveillance and Analysis: A Survey," *IEEE Access*, vol. 6, pp. 66816-66830, 2018.

[38] R. Mehran, B. E. Moore, M. Shah, "A Streakline Representation of Flow in Crowded Scenes," *European Conference on Computer Vision*, pp. 439-452, 2010.

[39] X. Zhang, D. Lin, J. Zheng, X. Tang, Y. Fang, H. Yu, "Detection of Salient Crowd Motion Based on Repulsive Force Network and Direction Entropy," *Entropy*, vol. 21, no. 6, pp. 608, 2019.

[40] R. Chaker, Z. Al Aghbari, I. N. Junejo, "Social Network Model for Crowd Anomaly Detection and Localization," *Pattern Recognition*, vol. 61, pp. 266-281, Jan. 2017.

[41] G. M. Basavaraj, A. Kusagur, "Optical and Streakline Flow Based Crowd Estimation for Surveillance System," *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*, pp. 414-416, 2016.

[42] D. Zhang, J. Xu, M. Sun, Z. Xiang, "High-density Crowd Behaviors Segmentation Based on Dynamical Systems," *Multimedia Systems*, vol. 23, no. 5, pp. 599-606, 2017.

[43] J. J. Monaghan, "Smoothed Particle Hydrodynamics and Its Diverse Applications." *Annual Review of Fluid Mechanics*, vol. 44, no. 44, pp. 323-346, 2011.

[44] H. Ullah, M. Uzair, M. Ullah, A. Khan, A. Ahmad, W. Khan, "Density Independent Hydrodynamics Model for Crowd Coherency Detection," *Neurocomputing*, vol. 242, pp. 28-39, 2017.

[45] B. K. P. Horn, B. G. Schunck, "Determining Optical Flow: A Retrospective," *Artificial Intelligence*, vol. 59, no. 1-2, pp. 81-87, 1993.

[46] X. Zhang, X. Shu, Z. He, "Crowd Panic State Detection using Rntropy of The Distribution of Enthalpy," *Physica A: Statistical Mechanics and its Applications*, vol. 525, pp. 935-945, 2019.

[47] D. Comaniciu, P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.

[48] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973.

[49] A. B. Chan, Z.-S. J. Liang, N. Vasconcelos, "Privacy Preserving Crowd Monitoring: Counting People Without People Models or Tracking," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-7, 2008.

[50] "Unusual Crowd Activity Dataset of University of Minnesota, <<http://mha.cs.umn.edu/Movies/Crowd-Activity-All.avi>>."

[51] W. Chu, H. Xue, C. Yao, D. Cai, "Sparse Coding Guided Spatiotemporal Feature Learning for Abnormal Event Detection in Large Videos." *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 246-255, 2019.

[52] X. Yan, Y. Ye, X. Qiu, H. Yu, Synergetic information bottleneck for joint multi-view and ensemble clustering, *Information Fusion*, vol. 56, pp. 15-27, 2020.



Xuguang Zhang received the B.S. degree in Electrical Technology from Northeast Normal University, and Ph.D. in Machinery and electronics engineering from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, China, in 2000 and 2008 respectively. He was a professor at the school of electrical engineering at the Yanshan University, China. He is currently a professor at the School of Communication Engineering, Hangzhou Dianzi University, China. His research interests include video and image processing, crowd behavior analysis, and human behavior understanding.



Dingxin Ma received the BS degree in communication engineering from Hangzhou Dianzi University, China, in 2018. Currently, he is pursuing the Master's degree in Hangzhou Dianzi University, China. His research interests is crowd video analysis.



Hui Yu is a Professor with the University of Portsmouth, UK. His research interests include vision, computer graphics and application of machine learning to above areas, particularly in human machine interaction, image processing and recognition, virtual/augmented reality, 3D reconstruction, robotics and geometric processing of human/facial performances. He serves as an Associate Editor of IEEE Transactions on Human-Machine Systems and the Neurocomputing journal.



Ya Huang has been a Senior Lecturer in Engineering Dynamics at the University of Portsmouth (UP) since 2009. He leads research into human responses to shock and vibration as part of the Bioneer Research Group and takes the initiative in human-centred autonomy research of future land and marine transports in the UP Future and Emerging Technology Theme and the Intelligent Transport Cluster. Before Portsmouth, he enjoyed 18-month working as part of the the Blast and Impact Dynamics Group at Sheffield University. Dr Huang obtained his doctorate degree in human responses to whole-body vibration in 2008 at the Institute of Sound and Vibration Research (ISVR), University of Southampton.



Peter Howell a Senior Lecturer in Computer Games Design in the School of Creative Technologies at the University of Portsmouth. He has design and development experience in the UK games industry, including in his role as Co-Designer at The Chinese Room. Peter completed his PhD in 2015, with his thesis focusing on Disruptive Game Design - an approach to game design that aims to subvert 'typical' game rules and associated player expectations of how games 'should' work.



Brett Stevens joined Portsmouth in 2001 and started teaching multimedia production, animation and virtual reality at Masters level. Since then, he has become the Director of Postgraduate Programmes for the School of Creative Technologies, overseeing all Masters curriculum development. His research has focused on Augmented Reality and Virtual Reality for many years, and he currently supervises PhD students in the areas of Virtual Reality, Games and Interactive interfaces.